La traducción imposible: poética de una voz.
Ensayo sonoro, ética del artificio e inteligencia artificial.
Tres giros en torno a la memoria de Mercedes Sosa**.**

Silvia Inés Vives

2025

Algunas palabras previas…

Esta obra no contiene mi voz. Pero todo en ella suena con la imposibilidad de decir que me habita. No se escuchan mis palabras, pero están ahí: en la grieta de la garganta, en el quiebre de una nota, en aquellos gestos que no tienen traducción.

Mi historia —como la de tantas— no fue dicha del todo.
No porque no existiera, sino porque no fue escuchada.

Trabajé con la voz de Mercedes Sosa no para imitarla, sino para dejar que dijera lo que no puedo decir. Ella canta desde un cuerpo que fue alguna vez desautorizado, exiliado, racializado, excluido, y en ese decir, mi silencio encuentra alguna forma.

En esa pieza habita mi voz —mi decir imposible— resonando en el mundo.

"Qué cosa fuera, corazón"
—Silvio Rodríguez,
interpretado por Mercedes Sosa

RESUMEN:

Este trabajo explora las tensiones entre la clonación vocal mediante inteligencia artificial y los significados poéticos, éticos y culturales encarnados en la voz humana. A partir de un corpus sonoro de Mercedes Sosa, se analiza el proceso técnico de extracción y resíntesis tímbrica, y se interroga qué puede (y qué no puede) ser replicado por la tecnología cuando se trata de una voz cargada de historia, identidad y afecto colectivo. Lejos de la mímesis perfecta, el experimento deviene una reflexión crítica sobre el artificio, la pérdida y el respeto por lo intrasferible.

OBJETIVOS:

Interrogar los alcances y límites de la clonación vocal mediante IA desde una perspectiva estética, técnica y ética.

Reflexionar sobre la voz como portadora de identidad, memoria y afecto colectivo, a partir del caso de Mercedes Sosa.

Articular herramientas técnicas con perspectivas del arte sonoro, la ética de la voz, como territorio de la memoria, la identidad afectiva individual y colectiva.

Palabras clave
arte sonoro, clonación vocal, Inteligencia artificial, ética de la voz

I.INTRODUCCIÓN

> *"Y los alemanes dijeron, ¿eso canta? Nunca van a entrar acá, porque yo no puedo traducir esas canciones. Entonces… los textos de los hermanos (..)*
> *Todas nuestras palabras de amor: chiquitita, queridita, negrita….*
>
> *No existe para ellos, en la traducción….*
> *Siempre es 'mi amor', nada más.*
>
> *Entonces… siempre pienso que…*
> *ahí, lo más difícil de todo,es traducir la poesía.*
> *En todos los idiomas del mundo cuesta esto.*
>
> *Pierde muchísimo.*
> *Pierde enormemente.*
> *Entonces, hay que hacer lo más parecido a lo que los poetas compusieron."*
>
> ***— Mercedes Sosa, Entrevista por O'Donnell,***
> ***Archivos O'Donnell, T1, C27: Especial Mercedes Sosa.***

Recientemente, el avance de las tecnologías de inteligencia artificial ha transformado radicalmente la relación entre arte, voz y máquina. Una de sus expresiones más inquietantes es la clonación vocal, proceso mediante el cual se replica el timbre de una voz humana para que diga —o cante— cualquier texto, incluso aquel que jamás emitió. Este fenómeno abre un campo fértil de experimentación, pero también de tensiones éticas, estéticas y culturales que invitan a repensar qué es lo que realmente encarna una voz.

El interés por este trabajo surgió a partir de una experiencia particular compartida en el marco de un seminario sobre IA[1]: la escucha de un chamamé de Tarragó Ros cantado originalmente por Violeta Pezzelatto y luego transformado tímbricamente, mediante IA, con la voz de un cantante de heavy metal. Esa operación técnica, en apariencia extraña, despertó en mí una pregunta urgente: ¿qué recupera la inteligencia artificial cuando copia una voz? ¿Qué se pierde en esa traslación?

A partir de esa inquietud, decidí trabajar con la voz de Mercedes Sosa, figura profundamente arraigada en nuestra cultura, cuya voz no solo canta, sino que transmite historia, memoria y afecto colectivo. La elección no fue arbitraria: se trata de una voz encarnada, que fuera alguna vez exiliada, desautorizada, excluida y, sin embargo, poderosa, sensible, política. Clonar esa voz no podía ser un mero ejercicio técnico. Debía ser también una pregunta, un gesto de respeto.

Este trabajo busca explorar qué se pierde cuando una inteligencia artificial intenta replicar una voz cargada de sentido vital. ¿Puede la tecnología imitar lo emocional, la historia corporal, la expresión singular que vibra en una voz como la de Mercedes? ¿Qué queda por fuera de la clonación tímbrica?

Lejos de una celebración de eficiencia técnica, este proyecto propone una escucha crítica y situada, donde la falla tecnológica revela lo inasible de lo humano. El recorrido integra tres dimensiones: un abordaje teórico que intenta pensar la voz y los alcances éticos del uso de la inteligencia artificial, una

[1] Seminario de actualización disciplinar I, dictado por el Doctor Juan Ramos. Maestría en Arte Sonoro, Universidad Nacional de Quilmes.

experimentación técnica con procesos de extracción y clonación vocal, y finalmente, la creación de una pieza sonora que actúe como "ensayo sonoro". [2]

No se trata de demostrar eficacia, sino de señalar un límite.
De hacer audible aquello que la IA no puede contener: el gesto, el afecto, la historia.

II. Primer Giro. Mercedes en la huella (Duelo en bucle)

2.1 DISEÑO EXPERIMENTAL

> *"Bueno a ninguno le gusta escucharse cantar.... Uno... es como.. por supuesto, una cosa tan íntima de uno... aunque parezca mentira, aunque esté ante una multitud. La emisión de la voz, y lo que siente realmente... aunque sea una multitud... que cuando se vuelve a escuchar, es desagradable. Ahí ha pasado la vida, la familia, las comidas, las bebidas... cuando hemos tomado con los amigos, a lo mejor fuera del canto, o cuando salimos de hacer la actuación. Entonces... de repente.., ud se escucha, (...)*
>
> *Y a veces, el canto... cuando ud lo vuelve a escuchar, ya no es lo mismo.*
> *Ha cambiado el tiempo, no?"*
>
> ***— Mercedes Sosa, Entrevista por O'Donnell,***
> ***Archivos O'Donnell, Temporada 1, capítulo 27: Especial Mercedes Sosa.***

En una entrevista con O'Donnell, Mercedes Sosa menciona un episodio donde necesitó escuchar una grabación de un show en vivo, en la que ella sentía que había desafinado. La cantora dice: "Nosotros no corremos carreras de caballos. Algunos artistas niegan que desafinan, se niegan al fracaso". Esa conciencia del cuerpo sonoro como territorio fallido y humano es clave para mí. A partir de la escucha de esa entrevista, surgió la decisión de trabajar con la voz de Mercedes Sosa, una figura cargada de sentido en nuestra cultura, cuya voz porta historia, identidad, emoción colectiva. La elección no fue arbitraria: Mercedes no sólo canta; enuncia, sostiene una ética de la voz que ella misma problematiza.

Esa dificultad que expresa en escucharse grabada, es también una conciencia del tiempo, del desfasaje de ese yo que ya no es. Escucharse es entonces enfrentarse al tiempo que pasó por la voz. Esta experiencia, entre íntima y abismal, resonó profundamente en mí y se volvió núcleo del experimento. ¿Qué fragmentos de esa historia —de ese cansancio, de ese afecto, de esa memoria— pueden sobrevivir al proceso de registro, reproducción, clonación artificial?

2.2. EXPLORACIÓN TÉCNICA

El trabajo de exploración técnica se centra en el entrenamiento de un modelo RVC a partir de un dataset vocal propio, construido con registros de Mercedes Sosa. El objetivo fue explorar qué puede —y qué no puede— replicar la inteligencia artificial cuando se enfrenta a una voz atravesada por historia, afecto y

[2] Este trabajo está compuesto, entonces, por tres giros: un primer giro (ensayo sonoro: la degradación en bucle); un segundo giro (pieza musical: el espejo roto); y un tercer giro, este propio texto.

resistencia. El resultado es un “ensayo” sonoro en forma de degradación progresiva, una trabajo que parte de una pregunta ética y estética: ¿qué se pierde cuando una IA intenta clonar una voz cargada de identidad y memoria colectiva?[3]

Inspirado en “I Am Sitting in a Room”, de Alvin Lucier —donde una frase grabada se reproduce y regraba hasta que solo queda el eco del espacio, la resonancia, la arquitectura—, este trabajo propone un gesto análogo, pero ya no en un espacio acústico, sino artificial.[4] Hacer pasar la voz a través de la máquina, una y otra vez, para que se degrade. Lo que se busca no es conservar, sino erosionar: recuperar la huella a través del desgaste, escuchar el hueco que queda a medida que ese espacio se vacía de identidad, afecto y sentido.

Para eso, se diseñó un bucle de iteraciones donde la voz sintética —“Mercedes IA”— fue retroalimentada al modelo en sucesivas rondas. Con cada vuelta, la voz se volvió más plana, más ajena, más sintética. El timbre se neutralizó, el ritmo se volvió rígido, y los signos emocionales —suspiros, titubeos, quiebres— terminaron por transformarse en ruido.

La voz de Mercedes Sosa no solo canta: enuncia, habla desde una historia y un cuerpo. Trabajar con ella desde la IA no es solo un experimento técnico, sino un intento por comprender aquello que siempre queda fuera del cálculo.

Frente a un contexto de plena aceleración tecnológica, este trabajo no busca la eficiencia técnica, sino detenerse a escuchar lo que se pierde.

[3] Esta pregunta deriva en otra, no menos necesaria: ¿es la inteligencia artificial la que clona? Es imprescindible desplegar en algún pensamiento que pueda desmontar la ilusión de transparencia tecnológica: no es el pincel el que pinta, sino quien lo sostiene. La herramienta no es inocente, ni autónoma. La IA, como toda tecnología, es una mediación cargada de decisiones invisibles: qué datos se utilizan, qué modelos se entrenan, qué voces se imitan y cuáles se omiten. La ética, entonces, no reside en el código, sino en el decir y en la escucha, en la relación que tejemos con aquello que decidimos replicar —o no. La pregunta no es si la inteligencia artificial puede clonar una voz, sino quién, para qué, desde dónde y con qué consecuencias lo hace. Como sostienen diversas perspectivas críticas sobre la tecnología, la propiedad ética no reside en la herramienta, sino en su uso. Y entonces sus aplicaciones no son éticamente neutras, ni transparentes: encarnan decisiones políticas, valores culturales y formas de habitar el mundo. "No es la flecha, sino el indio", dicen con delicada agudeza —y no sin ironía— los guitarristas de nuestra música popular cuando elogian sus guitarras. Clonar la voz de una artista como Mercedes Sosa no es solo un acto técnico, sino un gesto cargado de implicancias simbólicas, históricas y afectivas.

[4] Resignificar el trabajo de Alvin Lucier es, para mí, también un gesto de homenaje. Desde los años sesenta hasta hoy —con Duchamp, Warhol y tantas prácticas desde entonces— la cita, la copia, el ready-made, el remix, el cover, y ahora el dataset, se han vuelto parte constitutiva del lenguaje artístico contemporáneo. Repetir no es simplemente reproducir: es reescribir desde otro tiempo, desde otro cuerpo. En un presente fragmentado, trabajar con materiales ajenos es, muchas veces, una forma de recuperar astillas de sentido de una cultura que va quedando estallada, de recomponer lo que el tiempo (o la tecnología) ha dejado en ruinas.

2.3. DESCRIPCIÓN PROCESO DE TRABAJO

"Toba dueño como antes del bagre y la miel"
— Ramirez y Luna
interpretados por Mercedes Sosa

A continuación, se detalla el proceso de trabajo paso a paso, desde una perspectiva técnica, pero atravesada por una escucha atenta y una intención estética y ética que guió cada decisión.

PASO 1. Curaduría dataset.
Se elaboraron y curaron dos datasets diferenciados: uno de voz hablada y otro de voz cantada, ambos construidos a partir de registros de Mercedes Sosa. La selección priorizó materiales expresivos y con presencia corporal en la voz. En especial, interesó que el material combinara la voz hablada, la cantada, y los quiebres de emoción que emergen en ambas.

PASO 2. Entrenamiento del modelo.
Con estos materiales se entrenaron dos modelos independientes (uno por cada tipo de voz) utilizando la plataforma Weights. Un modelo Mercedes IA- Hablada, entrenado para "hablar" clonando el timbre vocal de la cantora, y un segundo modelo Mercedes IA - cantado

PASO 3. Elaboración del material:
A continuación, se abordó ese mismo corpus sonoro desde una lógica inversa a los estándares habituales: en lugar de buscar registros limpios o técnicamente estables, se seleccionaron fragmentos cargados de afecto: momentos donde la voz se quiebra, balbucea, respira, titubea o se interrumpe para hacer lugar a la emoción o al pensamiento.

También se aislaron palabras clave de fuerte carga simbólica y afectiva —"uno… es desagradable", se niegan al fracaso", "pierde enormemente", *"palabras de amor", "chiquitita", "negrita", "fuerza negra", "tierra", entre otras*. Esta selección pretende ubicar "signos sonoros" que cargan una intensidad afectiva y cultural que la IA no puede computar sin reducir. Mi intención es que lo intraducible ubicado allí se convierta en núcleo poético y político del trabajo.

PASO 4. Iteración.
El material expresivo elaborado en el paso anterior fue ingresado al modelo como input. El output resultante —la voz sintética generada por la IA— fue luego retroalimentado al modelo en nuevas rondas de inferencia. Este bucle de iteraciones produjo una degradación progresiva.

Se realizaron 17 iteraciones sucesivas del modelo, retroalimentando en cada caso el output como nuevo input. El criterio de finalización no fue predefinido numéricamente, sino determinado a partir de una escucha atenta del proceso. La serie se interrumpió en el punto donde la degradación fonética y tímbrica alcanzó una ruptura sensible del sentido: cuando las palabras comienzan a disolverse en articulaciones fallidas, y el contenido semántico se vuelve irreconocible.

Aunque técnicamente podrían haberse generado más iteraciones -y tal vez, podría expandirse el experimento aún más allá, el corte respondió a una búsqueda poética y afectiva. Ese umbral —donde la voz ya no dice — fue percibido como un momento de quiebre, una cesura de carga afectiva en si misma, en la que lo humano se torna inasible. Ese punto límite no es el agotamiento del modelo, sino una elección estética que marca el cierre del gesto. Una intención, un reconocimiento, un elemento que dice: esto, termina.

PASO 5. Integración.

Las sucesivas inferencias del modelo fueron ensambladas en un archivo final. No se incluyeron las 17 iteraciones completas: algunas fueron omitidas en la edición, siguiendo una lógica más poética que exhaustiva. Esta integración tampoco busca decirlo todo. Hay elecciones, ausencias, omisiones que componen el gesto.

Además, en el proceso de edición se superpusieron brevemente cada iteración con la siguiente, generando una imbricación. Por ejemplo, Mercedes AI1 y Mercedes AI2, coexisten por un breve instante. Mi intención es que este último gesto de imbricación no solo señale el momento de la repetición, sino que exponga también una marca de lo imposible, el proceso como su fractura, una mueca de quiebre en ese gesto de “duelo en bucle” frente a la degradación y la pérdida.

En este proceso, lo imperfecto se convirtió en el núcleo del trabajo. Ni el material ni el modelo fueron diseñados para optimizar resultados, sino para evidenciar la pérdida. El experimento no busca que la IA diga bien, ni que imite con fidelidad, sino que exponga aquello que no puede decir. No se entrena para ser más eficiente, sino para fallar: para fracasar en su intento de sintetizar lo humano.

2.4. ESCUCHA ACTIVA: ANÁLISIS COMPARATIVO DE LAS ITERACIONES

Luego de ensamblar la pieza, se abordó una escucha crítica y sensible del material y se identificaron las siguientes fases en el proceso:

Primera vuelta: Pérdida del espesor emocional (primera ronda de clonaciones)

A medida que avanza el proceso, puede percibirse como la voz pierde expresividad: el tono emocional se aplana, salvo en momentos de quiebre muy marcados. La IA tiende a neutralizar lo afectivo, lo singular, lo encarnado.

Segunda: Desfiguración fonética y tímbrica (segunda ronda de clonaciones)

En varias pruebas, la voz resultante adopta inflexiones ajenas, generando acentos no reconocibles —por momentos cercanos al francés, inglés u otras entonaciones extranjeras—. ¿Podríamos pensar que la IA, al buscar patrones universales, disuelve lo identitario?

Tercera: Descomposición del sentido (última ronda)

Los fonemas comienzan a deformarse progresivamente, hasta desdibujar las pocas palabras originales. El “sentido” semántico se pierde, y lo que queda es ritmo, ruido, articulación fallida.

III. Segundo giro: Mercedes AI, el espejo roto (Una voz que no fue)

3.1. CONTRAPUNTO TÉCNICO: EXPERIMENTO "EFICIENTE" COMO GRUPO DE CONTROL

> *"Entonces, de repente… ud se escucha… como cuando yo escuché, ahora, lo que hice en el Teatro Ópera, el segundo día… que… algo me decía a mí, que no estaba bien… pero, cuando lo volví a escuchar, en el DAT, estaba desafinado. (…) Y bueno, estaba desafinado… la entrada… Entonces… A nadie le gusta sentirse que, siempre, en esa frase, está desafinada… Estaba la voz indudablemente cansada y no llegaba a la exacta nota que tiene que llegar…*
>
> *Y hay artistas que niegan eso. Se niegan a que desafinan. Se niegan al fracaso. Niegan el fracaso, como si nosotros corriéramos una carrera de caballos… Nosotros no corremos carreras, contra nadie. Es la música que hacemos nosotros…. Es lo más importante que tiene un ser humano cuando hace música: cantar."*
>
> ***— Mercedes Sosa, Entrevistada por O'Donnell,***
> ***Archivos O'Donnell, Temporada 1, capítulo 27: Especial Mercedes Sosa.***

Para contrastar el experimento central —construido deliberadamente sobre la falla, la grieta y lo intransferible— se desarrolló una segunda prueba: una clonación vocal que intentó hacerse de modo "eficiente". Para ello, se entrenó un modelo RVC a partir de un dataset cantado de Mercedes Sosa, compuesto por registros técnicamente limpios, estables, de dicción clara y sin fallas emocionales ni tímbricas.

Como grupo de control, se utilizó este modelo para generar una inferencia sobre una versión de "Summertime" (G. Gershwin) cantada por Ella Fitzerlad, pieza perteneciente al repertorio tradicional del jazz, pero ajena al repertorio de Mercedes Sosa. El material fue editado: se extrajo y optimizó el stem vocal, se aplicó un proceso de dereverb (mvsep.com)[5], y se eliminaron artefactos generados durante el procesamiento (Izotope RX8), con el objetivo de lograr una mejor integración con el modelo. Durante la generación de la inferencia, se ajustaron el tono y la envolvente para optimizar la adaptación al modelo (weights.com).

La elección de esta pieza, ajena al repertorio original de Mercedes Sosa, responde a su alta exigencia técnica: fraseo enriquecido por interpretación típica de la música popular, afinación precisa, vibrato controlado, articulación clara y amplio rango dinámico. Por otro lado, se eligió por la compatibilidad tímbrica y de registro entre Ella Fitzgerald y Mercedes Sosa —ambas cantantes con gran riqueza tímbrica en los graves pero con extensión amplia hacia los agudos, y una canción con tonos largos de emisión sostenida —, lo que facilitó una transición más fluida en el proceso de inferencia.

A diferencia del experimento principal, centrado en la falla y la pérdida, aquí el modelo fue entrenado para "acertar", evaluando su capacidad para replicar con fidelidad los rasgos técnicos de una voz cantada. Esta voz "eficiente" intentó generar y observar un proceso "exitoso" desde lo técnico, que pudiera replicar con fidelidad el timbre original, sin desvíos significativos. En la escucha final de cada inferencia, se identificaron algunos artefactos que quedaron luego intencionalmente enmascarados por el acompañamiento instrumental.

[5] Se realizó la prueba con las plataformas moises.ar y mvsep, logrando mejores resultados con esta ultima

3.2. ELECCIÓN DEL MATERIAL

"Oh, your daddy's rich and your ma is good-lookin'
So hush, little baby, don't you cry.."
—Summertime de G. Gershwin – Letra Ira Gershwin
Interpretada por Ella Fitzgerald

La elección de Summertime como pieza para el experimento "eficiente" tampoco fue azarosa. Desde el otro lado del mundo, en otro territorio, desde otra historia, otra mujer -negra- canta. Ella Fitzgerald entona un lamento que es también una oda al verano: un verano idealizado, promesa de calma en medio de una historia marcada por la opresión y la desigualdad racial. Traducir Summertime —como concepto, como palabra, como clima emocional— supone un desafío similar al de traducir las canciones de Mercedes. Ambas voces están atravesadas por historias que no caben en los fonemas. Lo que en inglés canta la espera de un tiempo mejor, en Mercedes resuena como memoria de exilio, de lucha, de ternura herida.

3.3. EL ESPEJO ROTO, UNA VOZ QUE NO FUE.

"Pensé en prenderle fuego, pero temí que la combustión
de un libro infinito fuera semejantemente infinita y
sofocara de humo al planeta.."
— Jorge Luis Borges, El libro de arena

La segunda parte del trabajo propone un experimento en espejo. Si en el primer ensayo la inteligencia artificial es llevada al límite de su imposibilidad expresiva —el fracaso de clonar una voz cargada de historia y afecto—, aquí se la invita a hacer lo contrario: a triunfar técnicamente, a replicar de manera "correcta" una voz que nunca existió. El resultado es perturbador: una Mercedes Sosa que canta en inglés.

Esta elección no es casual ni inocente. Mercedes nunca cantó en inglés, no por incapacidad vocal, sino por convicción política, cultural y lingüística. Su arte fue un acto de enunciación situada, profundamente arraigada en el territorio, en lo popular, en lo latinoamericano. Su decisión de cantar en español y en lenguas originarias no era estética, sino ética. Por eso, hacerla cantar "Summertime" —pieza icónica del jazz anglosajón— no es solo una proeza técnica: es una intervención fantasmal.

Esta voz artificial compone un mundo que no fue, pero que ahora puede ser. El artificio genera una versión ficcional de una Mercedes que canta en inglés, con afinación precisa, vibrato controlado, fraseo impecable. Todo es correcto, todo es posible. Pero todo es falso.

Lo que se escucha no es una versión alternativa de lo real, sino una negación del acontecimiento histórico de esa voz. La IA, al borrar el límite de lo posible, también borra la historia.

Hay aquí una dimensión que roza lo psicótico: una escena en la que la realidad queda suspendida y reemplazada por una ilusión perfecta, donde la fragilidad de lo humano es borrada por el simulacro. No hay cuerpo, no hay historia, no hay límite. Lo que se produce no es evocación, sino impostura.

Este gesto artificial recuerda al universo de Borges. Cuenta el autor sobre el acto de contemplar "el Aleph", una experiencia abrumadora y desestabilizadora, que sugiere que la posesión de conocimiento o de posibilidades infinitas puede ser una carga insoportable. Así también esta voz infinita, sin errores, sin marcas, sin pasado, se vuelve una carga. Como en los libros con citas apócrifas que Borges dejaba entre líneas, esta obra genera una trampa del ser: una voz que no existió nunca, pero que ahora —en un archivo .wav— puede escucharse. Como esas citas inventadas que se vuelven verdad por repetición, esta voz borronea los límites entre lo real y lo imaginario.

El modelo "eficiente" no logra una réplica, sino una ficción de alta fidelidad: una versión que suena bien, pero que no dice nada. O peor: dice lo que no fue, lo que no podría haber sido, pero al decirlo lo niega. En esa negación de la pérdida, el gesto se vuelve exceso: un esfuerzo desbordado, un ejercicio de imposibilidad. El artificio intenta escuchar a Mercedes otra vez, tal vez solo por el placer de escucharla, o por la necesidad de escucharla, para escucharla un poco más. Y al hacerlo, la desarma.

Esta escena produce una forma particular de violencia simbólica.[6] Es un intento de recomponer el mundo estallado, pero con astillas reordenadas de manera arbitraria. Un mundo inventado, no solo improbable sino donde todo es posible, incluso a costa de negar lo frágil de la existencia. El resultado pone en crisis el criterio mismo de realidad: Mercedes canta en inglés. No como gesto de apertura, sino como simulacro. No hay memoria ni cuerpo en esa voz. Hay un eco sin huella.

El modelo eficiente logra lo imposible, y expone una grieta aún más honda: la pérdida de lo singular, de lo frágil, de lo finito, de lo fragmentado, de lo pequeño, de lo humano. En lugar de reconstruir el mundo que estalló —como el intento de rehacerlo con astillas—, lo reemplaza por un mundo paralelo, frío, sin historia, sin duelo. Y nos deja, otra vez, con una pregunta: ¿Qué pasaría si ardiera un libro infinito? ¿Qué pasaría si el mundo contara con esa biblioteca imposible donde todo puede ocurrir?

[6] Podría alguien preguntarse —con razón—, ¿desde qué lugar esta mujer latinoamericana pero bonaerense, blanca, digitalizada, nieta de europeos, toma la voz de Mercedes Sosa? ¿No es ese gesto, en sí, también una forma de apropiación, de borramiento, de violencia? No tengo una respuesta para eso. Pero puedo decir que este trabajo nace desde un profundo respeto, y desde una conmoción verdadera. La voz de Mercedes Sosa es, para mí, recuerdo de infancia, memoria de pueblo, afecto. Su voz en mi trabajo quiere ser -en lo que me alcanza-, un homenaje: imperfecto, pero guiado por el respeto y el cuidado.

3.4. DESCRIPCIÓN DEL PROCESO.

""Y a la hora del naufragio y la de la oscuridad,
alguien te rescatará"
— María Elena Walsh
Interpretada por Mercedes Sosa

PASO 1. Curaduría del dataset y entrenamiento del modelo
Para comenzar, recuperé el dataset de la voz cantada de Mercedes Sosa, elaborado previamente para el primer giro del proyecto. Este fue actualizado mediante una nueva generación. Para ello, se entrenó el modelo con fragmentos seleccionados de grabaciones de las canciones "Indio Toba" y "Zamba para olvidar" (citadas en la bibliografía). El stem vocal fue extraído utilizando la plataforma Weight, y posteriormente restaurado tímbricamente mediante herramientas digitales como dereverb (Weights), EQ, deesser, con el objetivo de optimizar y unificar tímbricamente los registros.

PASO 2. Inferencia: MercedesIA canta Ella
En esta etapa se explora la posibilidad —imposible— de una Mercedes Sosa cantando en inglés. A partir de fragmentos de Summertime interpretados por Ella Fitzgerald (en distintas versiones, algunas junto a Louis Armstrong), se generaron inferencias vocales que reproducen con notable precisión el timbre y la textura de la voz de Mercedes.

Para facilitar una integración más fluida con el modelo al realizar la inferencia, el tema fue transpuesto hasta seis semitonos hacia abajo. Una de las grabaciones seleccionadas incluía también la voz de Armstrong, lo que implicó un procesamiento adicional: primero se aisló y eliminó su línea vocal; luego, esos fragmentos fueron reemplazados por otras interpretaciones de Fitzgerald. Así se obtuvo una versión íntegramente cantada con su timbre, que fue posteriormente procesada por el modelo MercedesIA.

De este modo, la voz sintetizada de Mercedes interpretó la pieza conservando la cadencia de Ella y, de forma más sutil, por momentos con ciertos rastros del fraseo original de Armstrong. Durante todo este proceso —que implicó curaduría, separación de stems y ajustes técnicos— se buscó que la IA "hiciera bien su trabajo". Como resultado, la mayoría de los archivos suenan como Mercedes. Pero el efecto es inquietante: algo familiar, vuelto ajeno

PASO 3. Ensamblar el espejo roto
Finalmente, todo el material fue ensamblado. Summertime —antes interpretada por Ella y Louis— es ahora "reimaginada" en la voz de MercedesIA. A lo largo de la pieza se incorporan otras presencias vocales: fragmentos de Ella, de Louis, de Mercedes cantando su repertorio original, artefactos surgidos en el proceso, y versiones fantasmales de esas voces.

Estos materiales se suman en capas: aparecen voces desplazadas en su timbre o fraseo, pero también desplazadas tonalmente, alteradas en altura o ritmo, octavadas, transpuestas y desancladas de sus acompañamientos originales mediante técnicas de separación por stems. La intervención no se limita a lo vocal: también se aislaron y manipularon stems instrumentales que, una vez extraídas, fueron descompuestas y reorganizadas.

Progresivamente, los materiales se fragmentan o se ubican en posiciones anómalas. Ya no sostienen una canción: se convierten en materia disponible, inestable, capas flotantes que se ensamblan, se rozan, se contradicen. El resultado es un entramado sonoro donde voces e instrumentos coexisten, y cada elemento parece estar, deliberadamente, un poco fuera de lugar.

Podemos pensar que el gesto del Primer Giro —la degradación iterativa— aquí se transforma en una interpolación perfecta que, paradójicamente, revela su fractura. Lo técnico ya no fracasa por límite, sino por exceso: dice lo que no debía ser dicho.

Paso 5. Terminar de romper el espejo

Luego del ensamblaje final, la obra es intervenida para "romper el espejo" y hacer visible la fractura. Se aplican filtros visuales mediante herramientas como iZotope RX y el espectrograma de Adobe, que alteran el espectro de frecuencias de los distintos stems generando grietas, diagonales y “astillas”, ya sea por sustracción o incorporación de formantes y armónicos.

En este proceso final, se recupera un hilo que lo vincula con la exploración del primer giro. La mueca en la parte A (primer giro) —esa breve imbricación entre Mercedes 1 y 2 en el bucle, esa marca de lo imposible, doblez del tiempo donde asoman restos de tecnología en lo humano— reaparece en este segundo giro, ahora como grietas en B: se interviene para incorporar “fisuras” en el espectro, interrupción en las formantes, incongruencias en el patrón, fragmentación del tiempo por interpolación de brevísimos silencios que operan como cortes. Son en todo caso, tachaduras: señales de las astillas de un espejo roto. Elaboradas de forma artesanal, estas intervenciones conservan huellas de lo humano en la producción tecnologizada; rastros humanos en la máquina que, en definitiva, no es quien crea ni es tampoco nunca del todo transparente.

IV. REFLEXIONES: ÉTICA Y POÉTICA DE LO INTRADUCIBLE.

"Los nadies: los ningunos, los ninguneados…
que no hablan idiomas, sino dialectos."
— Eduardo Galeano, Los nadies

Mi interés en este trabajo fue que la clonación tímbrica no sea un ejercicio de mímesis, sino un modo de pensar la voz como lugar de identidad, afecto, memoria. La voz no es solo un timbre: es una biografía portada, una memoria colectiva, una política encarnada. En ese sentido, este trabajo es también una pregunta: ¿puede la inteligencia artificial recuperar lo inmesurable, eso que resuena más allá del texto en un texto?

Hay voces que no pueden simplemente ser replicadas. Voces que portan una densidad histórica, afectiva y política que excede cualquier sistema de síntesis. La voz de Mercedes Sosa es una de ellas. No solo por su timbre inconfundible, sino por lo que representa: una biografía situada, una historia de exclusión y resistencia, una ternura invencible que se filtra incluso en un silencio.

Esta imposibilidad no es un límite del lenguaje, sino una expresión de su profundidad: lo intraducible es identidad, memoria, experiencia vivida.

Clonar su voz, entonces, implica preguntarse qué ética guía esa operación. ¿Qué se imita cuando se imita una voz? ¿Qué entiende la IA del tono Quom en la canción Indio Toba? ¿Qué hace con el diminutivo que nombra el amor —"negrita", "chiquitita"— que, como dice Mercedes, no tiene traducción al alemán? ¿Cómo computa la IA la falla emocional del canto que se quiebra cuando canta "Como la cigarra"? ¿Cómo "escucha" y resintetiza esa tonada norteña, mestiza, atravesada por la historia, por una identidad individual y colectiva que dejó huella en una artista sensible? ¿Cómo honrar, y no traicionar, lo que esa voz encarna?

En este trabajo, lo intraducible no se concibe como un obstáculo, sino como brújula y faro: lo que guía la escucha, lo que orienta la práctica artística hacia una zona de cuidado, respeto y pregunta.

En ese sentido, pretende transformar una herramienta de clonación en un dispositivo de escucha ética. Usar una tecnología diseñada para imitar como modo de hacer audibles sus límites, y en ese hueco hacer que emerja lo humano —lo que no se puede clonar.

Lejos de la perfección técnica, esta pieza ensaya una ética de la escucha. En un mundo que corre hacia la eficiencia, propone una pausa: detener el tiempo para escuchar lo que se pierde.

Mercedes misma lo dice: "Traducir la poesía es lo más difícil del mundo. Se pierde muchísimo, se pierde enormemente". Tal vez, entonces, esta experiencia no sea sólo tecnológica, sino ética: un intento de respetar lo que no se puede clonar. Un intento por tensionar al límite la tecnología para recuperar lo humano.

Primer Giro

Mercedes ahí, en la huella

(Duelo en bucle)

Disponible en:

https://archive.org/details/primer-giro-mercedes-en-la-huella-duelo-en-bucle-v-3-09-06-25

Segundo Giro

Mercedes AI, el espejo roto.

(Una voz que no fue)

Disponible en:

https://archive.org/details/segundo-giro-mercedes-ai-el-espejo-roto-la-voz-que-no-fue

V. 4° movimiento: Salir del giro (Modelo para armar)

Es muy difícil saber de quién será el material que elaboré en este trabajo… Pero me animo a decir que este último movimiento, ~~tampoco~~ es mío. Dejo aquí sus materiales, para quienes quieran componer (o recomponer) esa voz.

Puntos de partida:

Mercedes Sosa - Dueño Antiguo de las Flechas
Mercedes Sosa interpretando Antiguos dueños de las flechas, junto con sus palabras finales, en un acústico en vivo grabado en Suiza en 1981.
Recuperado en Mayo de 2025 desde:
https://www.youtube.com/watch?v=AaHC4ay9FOU

Especial Mercedes Sosa - Archivos O'Donnell,
Temporada 1, capítulo 27: "Especial Mercedes Sosa".
Una entrevista realizada por Pacho O'Donnell en la que Mercedes reflexiona sobre su vida, su exilio, su voz y la música.
Recuperado en Mayo de 2025 de
https://www.youtube.com/watch?v=Fjx9FJ-nVdQ

Materiales elaborados y archivos:

Primer Giro (Duelo en Bucle)
Mercedes ahí, en la huella.
Degradación - iteraciones

- carpeta Iteraciones (archivo de inferencias)[7]
- dataset_mercedes_hablado.wav
- modelo_mercedes_hablado.rar[8]
- mercedesahi_boceto.wav
- voz_original_fragmento.wav

Disponible en:
https://drive.google.com/drive/folders/14PVf7O9v5CJ9-NHwXUEqVlql2OtcDwkQ?usp=drive_link

Segundo Giro (Una voz que no fue)
Mercedes AI, el espejo roto.
Contrapunto técnico.

- carpeta fantasmas (archivo de inferencias)[9]
- dataset_mercedes_cantado.wav
- modelo_mercedes_cantado.rar[10]
- mercedesIA_cantaElla_EJ2 (inferencia).wav
- voz_original_fragmento.wav[11]

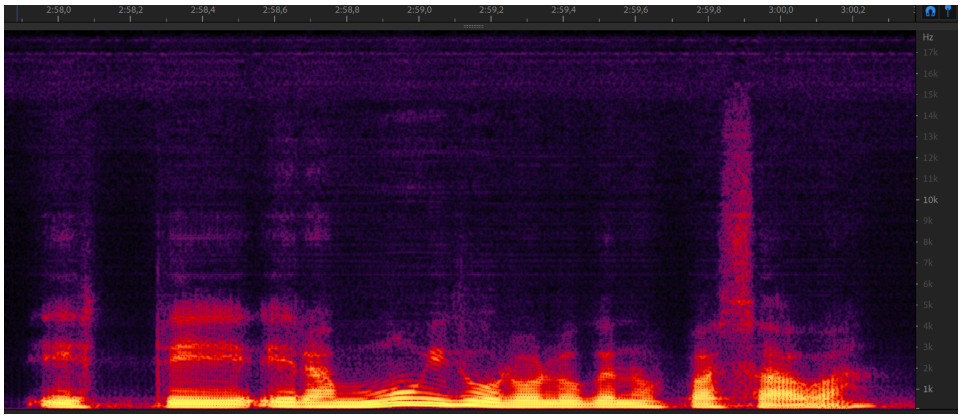
Disponible en:
https://drive.google.com/drive/folders/1uGSd_75d34iKK4Svw5Mc-vb3C6snFJVe?usp=drive_link

[7] Además del archivo final, con un primer boceto del trabajo, se incluye una carpeta con todas las iteraciones ordenadas, con el objetivo facilitar una escucha comparativa del proceso de degradación.
[8] Incluye 3 archivos: model.pth , model.index y model.json
[9] Además del archivo final, con un primer boceto del trabajo, se incluye una carpeta con inferencias de Summertime y repertorio de Mercedes en Inglés, con el objetivo facilitar una escucha comparativa del proceso.
[10] Idem 2
[11] Stem de Ella Fitzgerald - Summertime del Album Porgee and Bees, Recuperada en mayo 2025 en:
https://www.youtube.com/watch?v=2HJCN3upMHE

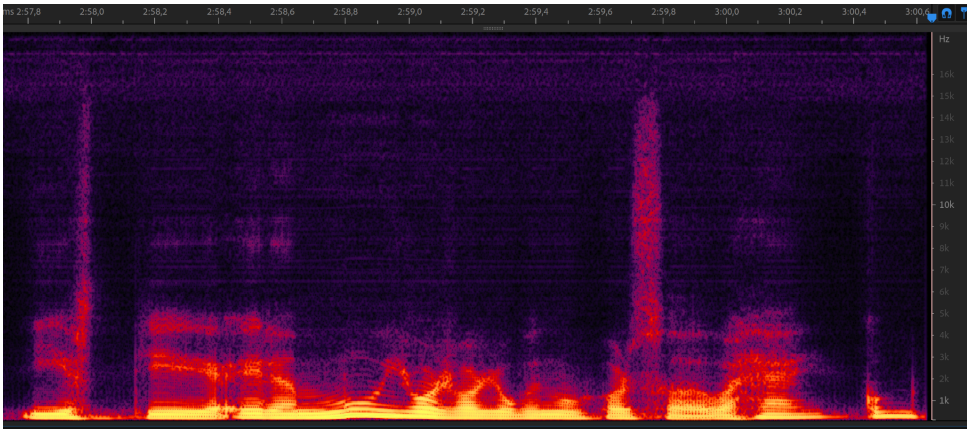
Nota final: Evidentemente, estos 62 elementos tampoco alcanzan para semejante título… pero he ahí otro humilde homenaje, memoria de mi humanidad.

Anexo. Registro imágenes

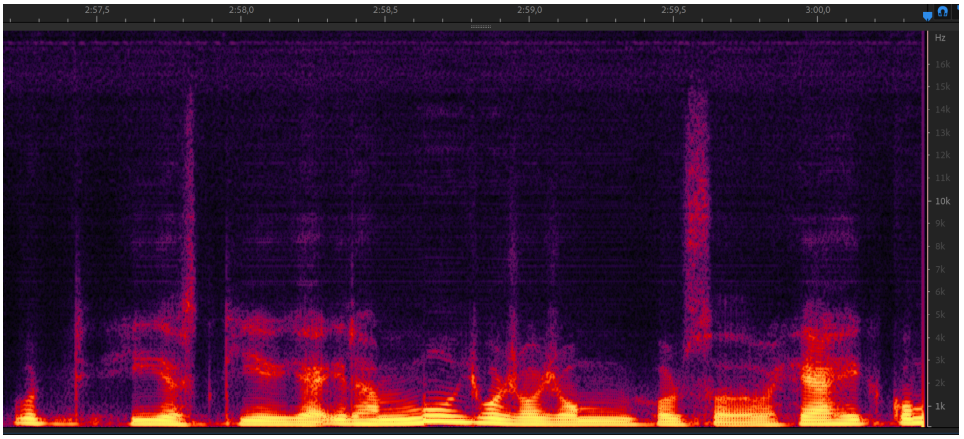
Primer Giro



Análisis espectral de fragmento de Primer Giro: entrevista, frases editadas sobre audio original
Audio de Mercedes sosa diciendo “Este.. era muy duro lo que nos pasaba”

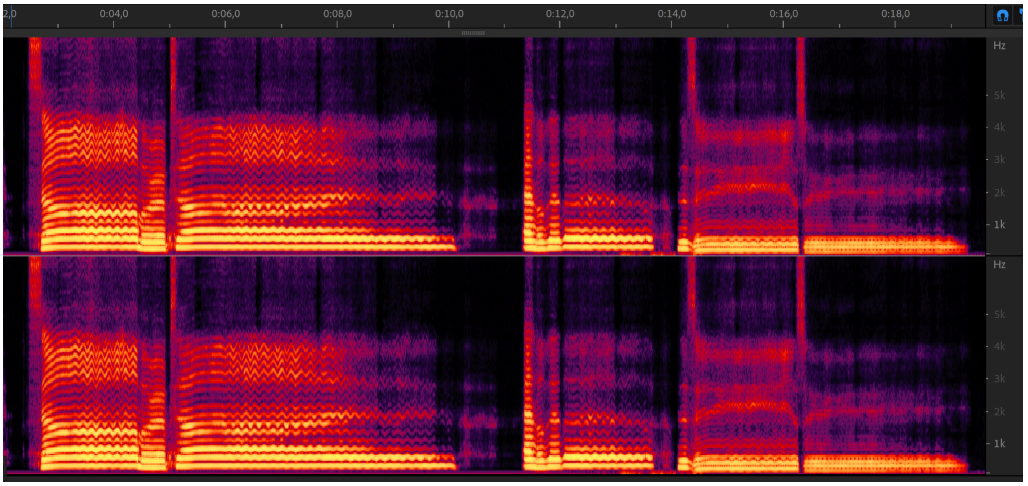


Análisis espectral de fragmento de Primer Giro: Inferencias
MercedesAI en la ronda de inferencias número 8 emitiendo la frase final “Este.. era muy duro lo que nos pasaba”

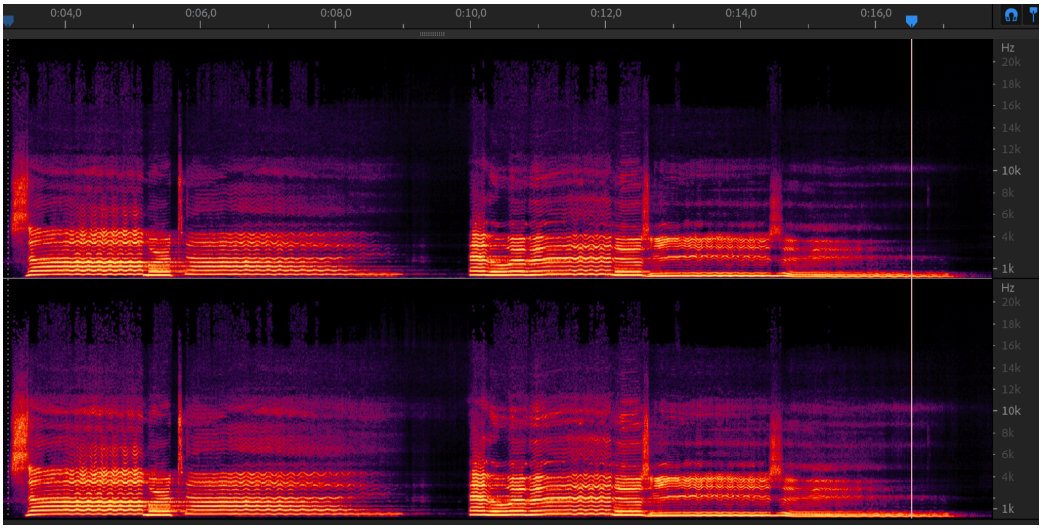


Análisis espectral de fragmento de Primer Giro: Inferencias
MercedesAI en la ronda de inferencias número 17 emitiendo la frase final “Este.. era muy duro lo que nos pasaba”

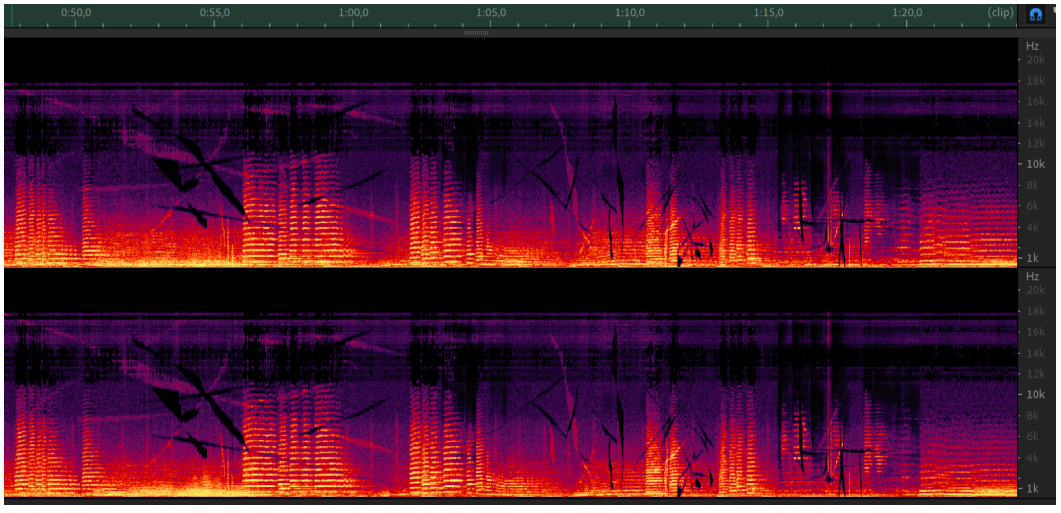
Segundo Giro



Análisis espectral de los materiales del Segundo Giro.
"Summertime and the living is easy" cantada por Mercedes AI



Análisis espectral de los materiales del Segundo Giro.
Análisis espectral del Steam Vocal de Ella FItzerald cantando "Summertime and the living is easy"



Análisis espectral de un fragmento de " Segundo Giro" - Tratamiento de filtros para "romper" el stem instrumental

Referencias

Barthes, R. (1972). *El grano de la voz*. En *Lo obvio y lo obtuso: Ensayos críticos III*. Buenos Aires: Ediciones Paidós. (Publicación original en francés: 1972).

Borges, J. L. (1949). *El Aleph*. Buenos Aires: Editorial Losada.

Borges, J. L. (1975). *El libro de arena*. Buenos Aires: Emecé Editores.

Cortázar, J. (1968). *62 / Modelo para armar*. Buenos Aires: Sudamericana.

Collins, N. (s.f.). *Composing for AI Voice Model Choir*. Durham University Music Department.

Galeano, E. (1998). Los nadies. En *El libro de los abrazos* (p. 170). Buenos Aires: Siglo XXI Editores.

Gershwin, G., & Gershwin, I. (1935). *Summertime* [Canción interpretada por Ella Fitzgerald]. En *Porgy and Bess* [Álbum]. Columbia Records.

Gershwin, G., & Gershwin, I. (1935). Summertime [Canción interpretada por Ella Fitzgerald en Vivo en Chicago 1958]. REmasterizado. Youtube https://www.youtube.com/watch?v=RHSJbN4e_sY

Kristeva, Julia, & Moi, T. (1986). World, dialogue and novel. En *The Kristeva Reader*. Columbia University Press.

Leguizamón, G., & Castilla, M. (1969). *Dueño antiguo de las flechas* [Canción interpretada por Mercedes Sosa]. En *Mercedes Sosa 30 años* [Álbum].

Leguizamón, G., & Castilla, M. (1969). *Dueño antiguo de las flechas* [Canción interpretada por Mercedes Sosa]. En *Acústico en Suiza 1980* [Presentación en vivo]. YouTube. https://www.youtube.com/watch?v=0wGPL32im0M

Lucier, A. (1981). *I Am Sitting in a Room* [Obra sonora].

Qian, K., Zhang, Y., Huang, M., Wu, Y., & Yang, D. (2022). *ContentVec: An improved self-supervised speech representation by disentangling speakers*. arXiv. https://arxiv.org/abs/2201.04200)

Ramos, J. M. (2025). *Introducción a la inteligencia artificial y redes neuronales* [Material de clase]. Seminario de Actualización en Sonido, Ciencias y Tecnología, Universidad Nacional de Quilmes recuperado en la página del Seminario de actualización disciplinar , Maestría en Artes Sonoros, Universidad Nacional de Quilmes -

Rodríguez, S. (1981). *La masa* [Canción interpretada por Mercedes Sosa]. En *Mercedes Sosa 30 años* [Álbum].

Segato, Rita. (2016). *La guerra contra las mujeres*. Editorial Prometeo / Traficantes de Sueños.

Schiller, F. (1785). An die Freude [Poema]. En L. van Beethoven (1824), Sinfonía n.º 9 en re menor, Op. 125 “Coral” [Obra musical], dir. H. von Karajan (1961–1962), Deutsche Grammophon citado por omisión en Vives, S. (2025). La traducción imposible [Persistencia sonora].

Sosa, Mercedes. (s./f.). Entrevistada por M. O’Donnell. En *Archivos O’Donnell*, Temporada 1, Capítulo 27: “Especial Mercedes Sosa” [Video]. YouTube. https://www.youtube.com/watch?v=Fjx9FJ-nVdQ

Toro, D., & Fontana, J. (2009). *Zamba para olvidar* [Canción interpretada por Mercedes Sosa con D. Torres]. En *Cantora* [Álbum]. Sony Music.

Walsh, Maria Elena (1972). *Como la cigarra* [Canción interpretada por Mercedes Sosa]. En *Mercedes Sosa 30 años* [Álbum].

Registro y derechos de autor
La presente obra ha sido registrada en la Dirección Nacional del Derecho de Autor (DNDA), República Argentina, bajo el número de expediente EX-2025-61695323-APN-DNDA#MJ (Primer Giro) / EX-2025-64766142- -APN-DNDA#MJ (Segundo Giro) / EX-2025-63302152- APN-DNDA#MJ (Tercer Giro). Los materiales sonoros asociados han sido publicados en la plataforma Archive.org y se encuentran protegidos bajo la licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0). Esta licencia permite su uso, distribución y adaptación con fines no comerciales, siempre que se otorgue el crédito correspondiente y se mantenga la misma licencia para las obras derivadas.